Appendix A

USEFUL RELATIONS AND DEFINITIONS FROM LINEAR ALGEBRA

A basic understanding of the fundamentals of linear algebra is crucial to our development of numerical methods and it is assumed that the reader is at least familar with this subject area. Given below is some notation and some of the important relations between matrices and vectors.

A.1 Notation

1. In the present context a vector is a vertical column or string. Thus

$$\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}$$

and its transpose $\vec{v}^T$ is the horizontal row

$$\vec{v}^T = [v_1, v_2, v_3, \ldots, v_m] \quad , \quad \vec{v} = [v_1, v_2, v_3, \ldots, v_m]^T$$

2. A general $m \times m$ matrix A can be written

$$A = (a_{ij}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ & & \ddots & \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix}$$

3. An alternative notation for A is

$$A = \left[\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m\right]$$

and its transpose A^T is

$$A^T = \begin{bmatrix} \vec{a}_1^T \\ \vec{a}_2^T \\ \vdots \\ \vec{a}_m^T \end{bmatrix}$$

4. The inverse of a matrix (if it exists) is written A^{-1} and has the property that $A^{-1}A = AA^{-1} = I$, where I is the identity matrix.

A.2 Definitions

1. A is *symmetric* if $A^T = A$.

2. A is *skew symmetric* if $A^T = -A$.

3. A is *diagonally dominant* if $a_{ii} \geq \sum_{j \neq i} |a_{ij}| \quad , \; i = 1, 2, \ldots, m$.

4. A is *orthogonal* if a_{ij} are real and $A^T A = AA^T = I$

5. $\bar{A}$ is the *complex conjugate* of A.

6. P is a *permutation* matrix if $P\vec{v}$ is a simple reordering of $\vec{v}$.

7. The *trace* of a matrix is $\sum_i a_{ii}$.

8. A is *normal* if $A^T A = AA^T$.

9. $\det[A]$ is the determinant of A.

10. A^H is the conjugate transpose of A, (Hermitian).

11. If

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

then

$$\det[A] = ad - bc$$

and

$$A^{-1} = \frac{1}{\det[A]} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

A.3 Algebra

We consider only square matrices of the same dimension.

1. A and B are equal if $a_{ij} = b_{ij}$ for all $i, j = 1, 2, \ldots, m$.

2. $A + (B + C) = (C + A) + B$, etc.

3. $sA = (sa_{ij})$ where s is a scalar.

4. In general $AB \neq BA$.

5. Transpose equalities:

$$\begin{aligned} (A+B)^T &= A^T + B^T \\ (A^T)^T &= A \\ (AB)^T &= B^T A^T \end{aligned}$$

6. Inverse equalities (if the inverse exists):

$$\begin{aligned} (A^{-1})^{-1} &= A \\ (AB)^{-1} &= B^{-1} A^{-1} \\ (A^T)^{-1} &= (A^{-1})^T \end{aligned}$$

7. Any A can be expressed as the sum of a symmetric and a skew symmetric matrix. Thus:

$$A = \frac{1}{2}\left(A + A^T\right) + \frac{1}{2}\left(A - A^T\right)$$

A.4 Eigensystems

1. The eigenvalue problem for a matrix A is defined as

$$A\vec{x} = \lambda\vec{x} \quad \text{or} \quad [A - \lambda I]\vec{x} = 0$$

and the generalized eigenvalue problem, including the matrix B, as

$$A\vec{x} = \lambda B\vec{x} \quad \text{or} \quad [A - \lambda B]\vec{x} = 0$$

2. If a square matrix with real elements is symmetric, its eigenvalues are all real. If it is asymmetric, they are all imaginary.

3. Gershgorin's theorem: The eigenvalues of a matrix lie in the complex plane in the union of circles having centers located by the diagonals with radii equal to the sum of the absolute values of the corresponding off-diagonal row elements.

4. In general, an $m \times m$ matrix A has $n_{\vec{x}}$ linearly independent eigenvectors with $n_{\vec{x}} \leq m$ and n_λ distinct eigenvalues (λ_i) with $n_\lambda \leq n_{\vec{x}} \leq m$.

5. A set of eigenvectors is said to be linearly independent if

$$a \cdot \vec{x}_m + b \cdot \vec{x}_n \neq \vec{x}_k \quad , \quad m \neq n \neq k$$

for any complex a and b and for all combinations of vectors in the set.

6. If A posseses m linearly independent eigenvectors then A is diagonalizable, i.e.,

$$X^{-1}AX = \Lambda$$

where X is a matrix whose columns are the eigenvectors,

$$X = \left[\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m\right]$$

and Λ is the diagonal matrix

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_m \end{bmatrix}$$

If A can be diagonalized, its eigenvectors completely span the space, and A is said to have a *complete* eigensystem.

7. If A has m distinct eigenvalues, then A is always diagonalizable, and with each distinct eigenvalue there is one associated eigenvector, and this eigenvector cannot be formed from a linear combination of any of the other eigenvectors.

8. In general, the eigenvalues of a matrix may not be distinct, in which case the possibility exists that it cannot be diagonalized. If the eigenvalues of a matrix are not distinct, but all of the eigenvectors are linearly independent, the matrix is said to be *derogatory*, but it can still be diagonalized.

9. If a matrix does not have a complete set of linearly independent eigenvectors, it cannot be diagonalized. The eigenvectors of such a matrix cannot span the space and the matrix is said to have a *defective* eigensystem.

10. Defective matrices cannot be diagonalized but they can still be put into a compact form by a similarity transform, S, such that

$$J = S^{-1}AS = \begin{bmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & J_k \end{bmatrix}$$

where there are k linearly independent eigenvectors and J_i is either a Jordan subblock or λ_i.

11. A Jordan submatrix has the form

$$J_i = \begin{bmatrix} \lambda_i & 1 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & \ddots & \vdots \\ 0 & 0 & \lambda_i & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 0 & \lambda_i \end{bmatrix}$$

12. Use of the transform S is known as putting A into its *Jordan Canonical form.* A repeated root in a Jordan block is referred to as a *defective* eigenvalue. For each Jordan submatrix with an eigenvalue λ_i of multiplicity r, there exists one eigenvector. The other $r-1$ vectors associated with this eigenvalue are referred to as *principal vectors.* The complete set of principal vectors and eigenvectors are all linearly independent.

13. Note that if P is the permutation matrix

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad , \quad P^T = P^{-1} = P$$

then

$$P^{-1} \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix} P = \begin{bmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{bmatrix}$$

14. Some of the Jordan subblocks may have the same eigenvalue. For example, the matrix

$$\begin{bmatrix} \begin{bmatrix} \lambda_1 & 1 & \\ & \lambda_1 & 1 \\ & & \lambda_1 \end{bmatrix} & & & & \\ & \lambda_1 & & & \\ & & \begin{bmatrix} \lambda_1 & 1 \\ & \lambda_1 \end{bmatrix} & & \\ & & & \begin{bmatrix} \lambda_2 & 1 \\ & \lambda_2 \end{bmatrix} & \\ & & & & \lambda_3 \end{bmatrix}$$

is both defective and derogatory, having:

- 9 eigenvalues
- 3 distinct eigenvalues
- 3 Jordan blocks
- 5 linearly independent eigenvectors
- 3 principal vectors with λ_1
- 1 principal vector with λ_2

A.5 Vector and Matrix Norms

1. The *spectral radius* of a matrix A is symbolized by $\sigma(A)$ such that

$$\sigma(A) = |\sigma_m|_{max}$$

where σ_m are the eigenvalues of the matrix A.

2. A p-norm of the vector $\vec{v}$ is defined as

$$||v||_p = \left(\frac{1}{M} \sum_{j=1}^{M} |v_j|^p \right)^{1/p}$$

3. A p-norm of a matrix A is defined as

$$||A||_p = \max_{x \neq 0} \frac{||Av||_p}{||v||_p}$$

4. Let A and B be square matrices of the same order. All matrix norms must have the properties

$$\begin{aligned}
||A|| &\geq 0 \text{ and } ||A|| = 0 \text{ means } [A] = 0 \\
||c \cdot A|| &= |c| \cdot ||A|| \\
||A + B|| &\leq ||A|| + ||B|| \\
||A \cdot B|| &\leq ||A|| \cdot ||B||
\end{aligned}$$

5. Special p-norms are

$$\begin{aligned}
||A||_1 &= \max_{j=1,\cdots,M} \textstyle\sum_{i=1}^{M} |a_{ij}| \quad \text{maximum column sum} \\
||A||_2 &= \sqrt{\sigma(\overline{A}^T \cdot A)} \\
||A||_\infty &= \max_{i=1,2,\cdots,M} \textstyle\sum_{j=1}^{M} |a_{ij}| \quad \text{maximum row sum}
\end{aligned}$$

   where $||A||_p$ is referred to as the L_p norm of $[A]$.

6. In general $\sigma(A)$ does not satisfy the conditions in 4, so in general $\sigma(A)$ is *not* a true norm.

7. When A is normal, $\sigma(A)$ *is* a true norm, in fact, in this case it is the L_2 norm.

8. The spectral radius of A, $\sigma(A)$, is the lower bound of all the norms of A.